
Poza widzeniem maszynowym. Jak zbudować nietrywialną maszynę percepcji

Joanna Żylińska

TEKSTY DRUGIE 2025, NR 1, S. 18–43

DOI: 10.18318/td.2025.1.2 | ORCID: 0009-0002-2829-2327

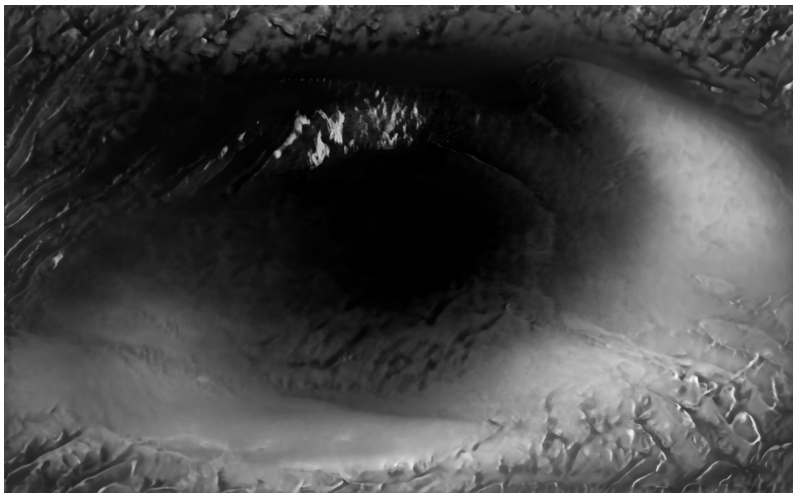
Źródło: Joanna Żylińska,
*Beyond Machine Vision.
How to Build a Non-Trivial
Perception Machine*,
„Transformations”, 2022
nr 36.

Jak widzieć lepiej niż ludzie

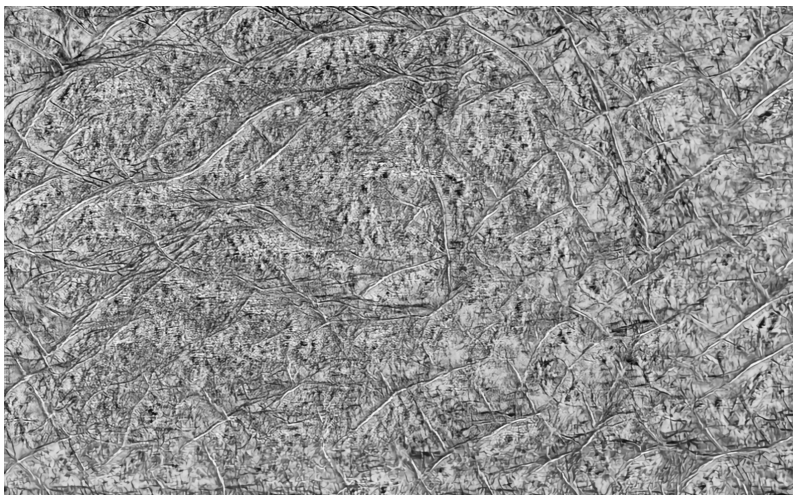
W niniejszym artykule omawiam zagadnienie sztucznej inteligencji i sztucznej kreatywności przez pryzmat widzenia maszynowego (il. 1 i 2). Uwypuklając machinacje towarzyszące wytwarzaniu koncepcji i praktyki widzenia u maszyn oraz u ludzi¹, rozpatruję w tekście wpływ modelu widzenia wypracowanego w informatyce na nasze wartości i instytucje społeczno-polityczne. Poza krytyczną analizą przedstawiam również twórczy eksperyment z zakresu inżynierii pojęciowej (*conceptual engineering*). Podejście to łączy pragmatyzm cechujący dziedzinę budowy maszyn z jej etymologicznym źródłosłowem, czyli łacińskim czasownikiem *ingeniare*, znaczącym tyle, co wymyślić, stworzyć

Joanna Żylińska – profesorka filozofii mediów i krytycznej praktyki cyfrowej w King's College London. Jest autorką licznych książek, w tym *AI Art. Machine Visions and Warped Dreams* (2020), *The End of Man. A Feminist Counterapocalypse* (2018) i *Nonhuman Photography* (2017). Współpracuje z innymi badaczami w ramach eksperymentalnych przedsięwzięć wydawniczych, takich jak *Photomediations* (2016). W działaniach artystycznych wykorzystuje rozmaite media posługujące się obrazami.

1 Ideę, że ludzkie widzenie jest nieodrodnie sztuczne, omawiałam już w *Nonhuman Photography*. Zob. J. Żylińska, *Nonhuman Photography*, MIT Press, Cambridge 2017. W rozdziale pierwszym tej książki analizuję człowieka jako „część złożonego asamblażu postrzegania, w którym z przyczyn funkcjonalnych, politycznych lub estetycznych schodzą się – i rozchodzą – różne organiczne i maszynowe podmioty”, s. 14.



Ilustracja 1 i z. Klatki z filmu *Neuromatic*, Joanna Żylińska (2020)



czy wykoncypować, a również zabawić się. Tytuł mojego artykułu nawiązuje do dwóch publikacji naukowych, napisanych w odstępnie dwudziestu lat przez badaczy, którzy wydatnie przyczynili się do debaty nad relacjami między ludźmi a maszynami. Odwołuję się do artykułu Heinza von Foerстера *Perception of the*

Future and the Future of Perception (Percepcja przyszłości i przyszłość percepcji) z roku 1971², w którym po raz pierwszy pojawiła się koncepcja „nietrywialnej maszyny”, oraz do tekstu Geralda M. Edelmana i George’a N. Reeke’a z roku 1990 zatytułowanego *Is It Possible to Construct a Perception Machine?* (Czy można skonstruować maszynę percepcji?)³. Rozważając te artykuły, spróbuję skonstruować koncepcyjne podwaliny teorii i praktyki percepcji maszynowej, a jednocześnie rozpatrzyć szerszy problem epistemicznej i rasowej (nie)sprawiedliwości w branży technicznej zajmującej się tym, aby maszyny „widziały”.

Na początku wypada wyjaśnić kluczowe dla mojego rozumowania pojęcia. Termin „widzenie maszynowe” odnosi się do gałęzi inżynierii systemów, która zajmuje się automatycznym wydobywaniem informacji z obrazów cyfrowych, co ma umożliwić maszynom wykonywanie zadań wymagających ludzkiego wzroku. Do takich zadań należą na przykład kontrola jakości oraz identyfikacja, lokalizowanie i pomiar, które realizowane są przez układy detekcji w samochodach autonomicznych, systemy ochrony czy też w badaniach kosmosu. Systemy widzenia maszynowego posługują się kamerami wyposażonymi w czujniki, sprzętem przetwarzającym oraz algorytmami oprogramowania. Oprogramowanie będące częścią tych układów wytwarzane jest przez pokrewną dziedzinę zwaną „widzeniem komputerowym”, która stanowi część rozległego pola sztucznej inteligencji (AI) i pod kątem teorii rozpatruje, w jaki sposób wydobywane są informacje z obrazów cyfrowych. Pomimo tej różnicy oba terminy – „widzenie maszynowe” i „widzenie komputerowe” – używane są czasem zamiennie. Celem widzenia maszynowego jest imitacja sposobu, w jaki ludzie widzą świat, lecz próba naśladownictwa zachodzi dopiero „po zdefiniowaniu procesów widzenia człowieka w kategoriach komputacyjnych”⁴. Kolejnym celem jest nauka tego, jak widzieć lepiej, czyli szybciej i wydajniej niż ludzie.

2 H. von Foerster, *Perception of the Future and the Future of Perception*, „Instructional Science” 1972, nr 1 (1).

3 G.M. Edelman, G.N. Reeke Jr., *Is It Possible to Construct a Perception Machine?*, „Proceedings of the American Philosophical Society” 1990, nr 134 (1).

4 Zob. L. Impett, *The Image-Theories behind Computer Vision*, referat przedstawiony na #DHNord2020 – *The Measurement of Images*, MESHs, zaktualizowany 29 września 2022 r., https://www.meshs.fr/page/the_image-theory_behind_computer_vision (20.12.2024). Moje ujęcie prezentowane w tym artykule jest pod pewnymi względami zbieżne ze stwierdzeniem Impetta: algorytmy widzenia komputerowego mają swoją ideologię, a także, co ważniejsze, swoją filozofię, chociaż w dominujących badaniach nad widzeniem komputerowym oba te aspekty zazwyczaj nie wbrzmiewają.

Teoretyczne przesłanki uczenia komputerów sztuki widzenia zostały sformułowane w opublikowanym w roku 1959 artykule autorstwa dwóch neurofizjologów, Davida Hubela i Torstena Wiesel, zatytułowanym *Receptive Fields of Single Neurons in the Cat's Striate Cortex* (Pola recepcyjne pojedynczych neuronów w korze prążkowanej kotów)⁵, który jest traktowany jako kamień węgielny widzenia komputerowego. Tekst oparty był na cyklu doświadczeń przeprowadzanych przez dwóch badaczy na „lekkich znieczulonych” kotach⁶. „Stosując ciągle dożylnie podawanie suksametonium, unieruchomiono” gałki oczne zwierząt⁷, aby zapobiec nieprzewidywalnym ruchom oczu. Kotom wyświetlano obrazy kropek i rozmaitych kształtów, aby zbadać aktywność ich mózgow i ustalić, które obszary kory skorelowane są z widzeniem. Okazało się, że projekcje świetlne różnych form nie wywołują żadnych znaczących zmian aktywności neuronowej, jednak przełom w doświadczeniu nastąpił, gdy u niektórych kotów zarejestrowano burzliwą aktywność neuronów na widok szczelin światła. Badacze zorientowali się, że kocie siatkówki reagują nie tyle na konkretny kształt, taki jak kropka czy kreska, ile na zmianę w intensywności światła na krawędzi obramowania slajdu. Hubel i Wiesel stwierdzili, że za aktywność tę odpowiadają neurony ułożone w kolumny w pierwszorzędowej korze wzrokowej (jak obecnie nazywamy ten obszar mózgu) i nadali im miano „komórek prostych”. Choć Hubel i Wiesel przeprowadzali swoje doświadczenia w warunkach laboratoryjnych i na unieruchomionych obiektach niebędących ludźmi, dały one początek jednemu z podstawowych założeń (a faktycznie mitów) widzenia komputerowego, a mianowicie pogładowi, że proces widzenia jest wielowarstwowy i hierarchiczny, że można wyodrębnić istotę widzenia u różnych zwierząt (w tym zwierząt ludzkiej odmiany), że mechanizm postrzegania krawędzi stanowi sedno widzenia, że proces ten jest fizjologiczny i niezależny od tego, co się widzi, i że maszyny można nauczyć widzieć „jak ludzie” przez odtworzenie procesu postrzegania wzorców (*pattern perception*) na poziomie pikseli.

Te formalistyczne parametry utrwalił dodatkowo podręcznik widzenia komputerowego z roku 1982 autorstwa Davida Marra, naukowca z Massachusetts Institute of Technology, zatytułowany po prostu *Vision* (Widzenie). Marr twierdził w nim, że widzenie stanowią przede wszystkim zjawisko z zakresu

5 D.H. Hubel, T.N. Wiesel, *Receptive Fields of Single Neurons in the Cat's Striate Cortex*, „Journal of Physiology” 1959, nr 148.

6 Tamże, s. 589.

7 Tamże, s. 575.

przetwarzania informacji⁸. Według tych założeń widzeniem zawiaduje mechanizm obliczeniowy, przebiegający od podstawowych procesów widzenia, takich jak wykrywanie krawędzi czy widzenie obuoczne, ku wyższym poziomom analizy obwodów neuronalnych rzekomo umożliwiającym widzenie, aż po cały mózg. Takie właśnie pojmowanie „widzenia komputerowego” i jego inżynierskiego odpowiednika, „widzenia maszynowego”, chcę poddać krytycznemu oglądowi, a następnie rozpatrzyć, jak maszyny mogą widzieć, a właściwie czy w ogóle mogą widzieć. Interesuje mnie zatem również, co oznacza dla ludzi to, że zdolność maszyn do rozróżniania obiektów w świecie na podstawie światła odbijanego przez nie i przekazywanego w formie danych do procesorów tych maszyn klasyfikuje się jako widzenie. Krytyczna analiza pozwoli mi rozważyć możliwość skonstruowania maszyny percepcji według założeń odwołujących się do reguł nakreślonych przez biologię i informatykę, ale też wykraczających poza nie.

Od widzenia maszynowego do percepcji maszynowej

Zwrot od widzenia ku percepcji stanowiący część tego eksperymentu nie jest całkowicie moim pomysłem. Otóż w ostatnich latach wielu badaczy AI zarzuciło ścisły okulocentryzm widzenia maszynowego, rozciągając swoje badania nad maszynowymi operacjami przechwytywania danych na inne zmysły, takie jak słuch, dotyk i węch. Firma Google, która do nieruchomych obrazów dorzuca obecnie inne źródła danych, takie jak dźwięki, muzykę i wideo, posługuje się już terminem „percepcja maszynowa” zamiast „widzenia maszynowego”. Nadal jednak operacje percepcyjne zasadzają się w niej na rozpoznawaniu przedmiotów przy użyciu algorytmów szkolonych na wielkich, częściowo oznaczonych zbiorach danych z wykorzystaniem klastrów przetwarzania równoległego. Stosowane przez Google pojęcie percepcji maszynowej obejmuje wprawdzie szerszy zbiór danych zmysłowych, lecz zachowuje ideę z góry określonych przedmiotów, które mają być zmysłowo odczuwane, wraz z ich właściwymi kategoriami⁹. Ja natomiast proponuję tutaj pojemniejszą koncepcję percepcji maszynowej niż ta przedstawiona przez Google, chcąc naświetlić jedną z głównych białych plam współczesnego widzenia komputerowego, a mianowicie to, że nie wyjaśnia ono, jak

8 Zob. D. Marr, *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*, W.H. Freeman, San Francisco 1982, s. 3-5.

9 Zob. <https://research.google/pubs/?area=machine-perception> (31.01.2025).

właściwie działają nasze mózgi i jak przebiega przekazywanie i przetwarzanie bodźców siatkówkowych w obwodach neuronalnych mózgu – czyli proces, w którym powstają nasze odczucia zmysłowe, a także świadomość tych odczuć (rozpoznanych w formie obrazów, które „widzimy”). Koncepcja ta stawia również pod znakiem zapytania założenie, że naczelnym organem postrzegania jest oddzielna fizjologiczna jednostka zwana mózgiem, a zatem kwestionuje modelowanie widzenia maszynowego na podstawie procesów korowych ludzkiego mózgu. W artykule z roku 1959 zatytułowanym *What the Frog's Eye Tells the Frog's Brain* (Co oko żaby mówi mózgowi żaby) kognitywista J.Y. Lettvin¹⁰ i współpracownicy wykazali, że aktywność percepcyjna, która, jak zakładano, przebiega w mózgu w efekcie stymulacji siatkówki przez światło, faktycznie zaczyna się już w oku. Gdzie dokładnie zachodzą procesy postrzegania i jak konkretnie działa ich podstawowy mechanizm – nadal nie wiadomo (i to zarówno u żab, jak i u ludzi). A zatem koncepcja percepcji maszynowej komplikuje model przedstawiający widzenie jako proces zasadniczo bazujący na reprezentacjach oraz podsuwa nowy sposób rozumienia, co właściwie miałyby znaczyć, że maszyny widzą. Koncepcja ta zakłada również, że postrzeganie odbywa się w świecie w takim samym stopniu co w oku czy mózgu. W obliczu tego przesunięcia akcentów potrzeba nam bardziej umiejscowionego, ucieleśnionego i dynamicznego rozumienia, jak komputery (a także ludzie) widzą świat – i jak w nim działają.

Postrzeganie przyszłości

Posługuję się tu perspektywą inżynierii pojęciowej w sposób odbiegający nieco od najnowszych zastosowań tego terminu w filozofii analitycznej, w której służy on jako bardziej programowe i mniej żartobliwe narzędzie myślowe. Na przykład David Chalmers definiuje inżynierię pojęciową jako „proces projektowania, wdrażania i oceny pojęć”¹¹. Uznaje on, że doprecyzowanie, a zatem ustabilizowanie pojęć jest co najmniej tak samo ważne jak wymyślanie ich. Dla mnie zaś inżynieria pojęciowa jest czymś na kształt cyborgicznego briko-lażu: jest mniej trwała, mniej zaprzątnięta obwarowywaniem się, natomiast

10 J.Y. Lettvin i in., *What the Frog's Eye Tells the Frog's Brain*, „Proceedings of the Institute of Electrical and Electronic Engineers” 1959, nr 47.

11 D. Chalmers, *What is Conceptual Engineering and What Should it Be?*, „Inquiry. An interdisciplinary journal of philosophy”, 16 września 2020, s. 2 (numer strony odnosi się do przedruku dostępnego na stronie internetowej autora).

o wiele bardziej przewrotna. Inżynieria pojęciowa jako podejście łączy w sobie dwie orientacje wyróżnione przez Chalmersa, a mianowicie „inżynierię *de novo*” (od nowa) oraz „re-inżynierię” (*re-engineering*), przygląda się wzorcom i podbudowom uznanych idei i tekstów, aby zmontować nowe, choćby tylko tymczasowe, konstrukcje, szukając przy tym wyjść awaryjnych.

Jak już wspomniałam, sięgnięcie po dwie publikacje – von Foerстера oraz Edelmana i Reeke’a – jest w tym kontekście zasadne, gdyż autorzy ci walnie przyczynili się do przededefiniowania fizjologicznych zjawisk takich jak postrzeganie i widzenie w kategoriach programowalności. Ich badania przekroczyły tradycyjne granice dyscyplin naukowych w poszukiwaniu nowych sposobów pojmowania ludzi i maszyn, a nawet ludzi jak o maszyn. Von Foerster, austriacki fizyk i inżynier, żywo interesował się biologią. Większość życia zawodowego spędził w USA i należał do grona najważniejszych uczestników konferencji w Fundacji Macy’ego – cyklu spotkań organizowanych w latach 1940-1961, których efektem było wyłonienie się cybernetyki jako transdyscyplinarnej dziedziny badań nad układami biologicznymi i mechanicznymi. Von Foerster był płodnym pisarzem, przy czym często wychodził poza obszar swojej dyscypliny, a nawet nauk ścisłych jako takich, aby przedstawiać ogólniejsze spojrzenie na społeczeństwo i świat. Skłonność tę świetnie obrazuje jego tekst *Perception of the Future and the Future of Perception* (Percepcja przyszłości i przyszłość percepcji), który przedstawił po raz pierwszy na otwarciu Dorocznej Konferencji Problemów Światowych na Uniwersytecie Kolorado w Boulder w roku 1971.

Na początku von Foerster przywołuje teorię systemów będącą sposobem myślenia o zmianie społeczno-kulturowej, a jednocześnie wskazuje na ograniczenia statycznego modelu systemu, w którym zmianę uważa się za aberrację, konieczną do usunięcia, a nie za zachętę do przekształcenia systemu. Biorąc pod uwagę mechanizm działania systemu, von Foerster przechodzi do poszukiwania w nim jakiejś możliwości urzeczywistnienia ludzkiej wolności i sprawczości, a w tym celu zwraca się właśnie ku postrzeganiu. Konkretnie zaś radzi: „jeśli pragniemy być podmiotem, a nie przedmiotem, to to, co widzimy w danej chwili, czyli nasza percepcja, musi wybiegać w przód, a nie spoglądać wstecz”¹². Innymi słowy, von Foerster nakłania, abysmy nie polegali na ustalonych wzorcach widzenia rzeczy, ale podążali niespodziewanymi ścieżkami i nowymi szlakami, ucząc się od dzieci. Na razie brzmi to dość idyllicznie; von Foerster nie jest oczywiście pierwszym myślicielem

12 H. von Foerster, *Perception of the Future and the Future of Perception*, s. 31.

przywołującym ideę pierwotnego, nieskażonego spojrzenia, które rzekomo daje lepszy wgląd w prawdę tego świata: przed nim czynili to filozofowie oświeceniowi nawołujący do powrotu do „stanu natury”, poeci doby romantyzmu i surrealiści. Jednak von Foerster ma też do powiedzenia kilka interesujących rzeczy o współczesnej atrofii percepcji, za którą wini utowarowienie informacji oraz „system oświaty, który myli proces tworzenia nowych procesów z rozdzielaniem dóbr zwanych «wiedzą»”¹³. Jako lek na tę bolączkę von Foerster wprowadza pojęcie „nietrywialnej maszyny”, w którym „maszyna” „odnosi się do jasno określonych funkcjonalnych właściwości pewnej abstrakcyjnej jednostki, a nie do zespołu trybików, przycisków i wajch, chociaż takie zespoły mogą stanowić ucieleśnienie owych abstrakcyjnych jednostek funkcjonalnych”¹⁴.

Podczas gdy w „maszynie trywialnej” „efekt wyjściowy” pozostaje w prostej relacji jeden-do-jednego z „danymi wejściowymi”, a zatem wytwarzane przez nią skutki są, przynajmniej teoretycznie, przewidywalne i zawsze takie same, w maszynie nietrywialnej „jej poprzednie działania warunkują jej bieżące reakcje”¹⁵. W obu tych systemach obowiązuje zasada determinizmu, lecz drugi z nich jest nieprzewidywalny, gdyż wynik zmienia się w zależności od przebiegu wcześniejszych cykli działania. Von Foerster ceni sobie solidną przewidywalność rozmaitych trywialnych systemów, takich jak tostery czy samochody, ale niepokoi go zastosowanie trywializacji w innych systemach. W tym właśnie punkcie jego argumentacja i koncepcja maszyny stają się szczególnie interesujące. Badacz ilustruje swój wywód przykładem studenta, w którym widzi potencjalnie nietrywialną maszynę ulegającą jednak zupełnej „trywializacji” po włączeniu się w maszynę wyższego rzędu, jaką jest uniwersytet z jego przewidywalnym nauczaniem i odtwórczym testowaniem: „Stuprocentowy wynik na teście wskazuje na stuprocentową trywializację: student ten jest całkowicie przewidywalny, a zatem można przyjąć go do społeczeństwa. Nie przysporzy on ani zaskoczeń, ani kłopotów”¹⁶. Choć człowiek (tj. student) uważany jest za maszynę, to w procesie konstruowania w stosunkowo młodym wieku operacje wyższego rzędu pętli rekursywnych umożliwiają kreatywność w ramach systemu złożonego z innych maszyn

13 Tamże, s. 33.

14 Tamże, s. 40.

15 Tamże.

16 Tamże, s. 41.

(wykładowców, form wiedzy i sposobów jej przekazywania, pedagogii, budynków, infrastruktur) – pod warunkiem jednak, że ludzkie podmioty zrealizują swój potencjał, czyli gdy przyjmą, że nieobliczalność jest częścią systemowej racjonalności.

Wydawać by się mogło, że wykładowcy, czy w rzeczy samej ludzie w wielkości, tego właśnie pragną i że chętnie skorzystaliby z wolności w dostępnym im, systemowo zakreślonym, lecz ostatecznie nieustalonym zakresie. A jednak takie akty korzystania z wolności, czyli próby przekierowania biegu działań systemu, są raczej rzadkie, co potwierdzają zarówno von Foerster (który niechęć tę uważa za rodzaj przypadłości), jak i filozof Vilém Flusser (którego zdaniem zawsze działamy w warunkach systemowej przewidywalności, czy to na poziomie maszyny, społeczeństwa, czy nieuchronnie entropicznego wszechświata¹⁷). Koniec końców w konkluzji artykułu von Foerster w całkiem humanistycznym duchu apeluje, aby dostrzegać i cenić „wichrzycieli”, których, jak mówi, poznamy „po akcie stworzenia: «Niech stanie się widzenie: i stało się światło»”¹⁸. Co interesujące, to postrzeżenie ma umożliwiać nam (jako osobom wykładającym i oceniającym studentów) wprowadzenie czegoś nowego w proces wyjście-wejście, a zatem wykreowanie pewnej systemowej sposobności. To zaś mogłoby stanowić pierwszy krok w próbie zbudowania nietrywialnej maszyny percepcji.

Ale jak właściwie mielibyśmy sprawić, aby studenci (czy w gruncie rzeczy ktokolwiek) widzieli lepiej, nie wiedząc z góry, co mieliby naszym zdaniem zobaczyć, jakie ramy ukierunkowują nasze patrzenie i jakie są naszym celem? Czyż nie grozi nam to, że apel von Foerstera wybrzmi banalnie czy nomen omen trywialnie: w najlepszym razie mgliście humanistycznie, w najgorszym zaś jak przepis na zatracenie przez oświatowych libertarian wszelkiej fachowości i troski wypracowanej w systemie edukacji? Von Foerster, nazywany człowiekiem renesansu, jest jednym z wielu naukowców (typu męskiego), którzy wygłaszając sądy na temat szerszych kwestii społecznych, posiłkują się pobieżnie modelami naukowymi i protokołami badawczymi, aby tłumaczyć świat, a jednocześnie wydają się nie dostrzegać kulturowych

17 Zob. V. Flusser, *Ku filozofii fotografii*, przeł. J. Maniecki, przedm. P. Zawojski, Aletheia, Warszawa 2015, s. 65-71; tenże, *Into the Universe of Technical Images*, przeł. N.A. Roth, University of Minnesota Press, Minneapolis 2011, s. 19, 89-92. [Fragment drugiej pozycji ukazał się w polskim przekładzie Andrzeja Gwóźdź (zob. V. Flusser, *Ku uniwersum obrazów technicznych*, w: *Po kinie? Audiowizualność w epoce przekazników elektronicznych*, red. A. Gwóźdź, Universitas, Kraków 1994), lecz publikacja ta nie obejmuje przywoływanego tu ustępu – przyp. tłum.]

18 H. von Foerster, *Perception of the Future and the Future of Perception*, s. 43.

aspektów opisywanych przez nich systemów, w tym terminologicznego (i ideologicznego) uwarunkowania swoich pojęć. (Oczywiście wszystkie założenia koncepcyjne są ideologicznie uwarunkowane, ale problemem, na który tu wskazuję, jest nieświadomość lub maskowanie tego zjawiska w połączeniu z przypisywaniem „ideologiczności” swoim intelektualnym przeciwnikom).

Taki właśnie model „okrojonej racjonalności” działa w Dolinie Krzemowej, siedlisku myśli cybernetycznej i postcybernetycznej. Ma się on również świetnie w wielu dziedzinach naukowych, takich jak filozofia analityczna, językoznawstwo, biologia, psychologia kognitywna i neuronauka, tworzących intelektualną podbudowę sztucznej inteligencji i często korzystających z hojnego finansowania udzielanego przez sektory technologii i wojskowości. W reakcji na wszechobecne przejawy okrojonej racjonalności w naszym – coraz bardziej zautomatyzowanym i rządzonego przez dane – społeczeństwie Luciana Parisi zaproponowała hipotezę obcego (*Alien Hypothesis*). Jest to koncepcja, która przyjmuje, że „abdukcyjne, konstrukcjonistyczne, eksperymentalne wyobrażenie działania logiki” stanowi część „spekulatywnego obrazu obcego podmiotu sztucznej inteligencji”¹⁹. W idei Parisi interesujące jest zwłaszcza to, że wyłania się ona z logicznych parametrów myśli cybernetycznej. Jej hipoteza ma być alternatywą zarówno wobec hipotezy cybernetycznej, zakładającej, że jedyną ucieczką od ograniczeń systemu obliczeniowego jest zaszczytowanie się w strefach nieprzejrzystości i niewidoczności, co łączy się z całkowitą negacją logiki systemu, jak i wobec hipotezy akcelerationistycznej, która nawołuje do demontażu władzy systemu od środka na drodze wyczerpania jego operacji. Parisi natomiast proponuje włączenie się w działania systemowe na ich własnych warunkach logicznych z zamysłem odnalezienia „obcej przestrzeni rozumowania”²⁰ wewnątrz systemu. Mając na uwadze implikacje spostrzeżeń Gregory’ego Chaitina, odkrywcy liczb nieobliczalnych w systemach obliczeniowych, Parisi dowodzi, że przestrzeń ta kryje się pomiędzy serwowym i zaprogramowanym i zapętłonym serwowym systemem, która wytwarza złudzenie, że inne opcje nie istnieją. Rozwiązanie przedstawione przez Parisi leży w przededefiniowaniu medium myśli wychodzącym poza jej automatyzację i instrumentalizację w dzisiejszych systemach maszynowych. Moje rozwiązanie natomiast polega na cofnięciu się o krok w naszym spektrum poznawczo-zmysłowym, tak aby głównym trybem wchodzenia

19 L. Parisi, *The Alien Subject of AI*, „Subjectivity” 2019, nr 12, s. 32.

20 Tamże, s. 30.

w relacje ze światem, dzięki któremu można w ogóle przystąpić do próby demontażu działań systemu, była percepcja, a nie myśl²¹.

Widzenie poza mózgiem

A zatem czy można skonstruować maszynę percepcji? Jak już wspomniałam, pytanie to postawili najpierw biolog i laureat Nagrody Nobla Gerald M. Edelman i jego współpracownik George N. Reeke Jr. w artykule z roku 1990. Publikacja Edelmana i Reeke'a była częścią szerszego przedsięwzięcia poświęconego lekceważeniu odkryć biologii ewolucyjnej w badaniach nad sztuczną inteligencją. W tekście napisanym dwa lata wcześniej naukowcy ci zganili badaczy AI za nadmierne przywiązanie do

epistemologicznych założeń wyprowadzonych z jednej strony z wywodów Alana Turinga i Alonza Churcha o uniwersalnej zdolności komputerów do rozwiązywania problemów (wskazujących, że mózg można pojmować jako komputer), a z drugiej strony z redukcjonizmu biologii molekularnej (wskazującej, że mózg można pojmować jako zbiór jednostek wymieniających sygnały chemiczne)²².

Edelman i Reeke mieli jeszcze inne, raczej krytyczne uwagi o przyjęciu „informatyki opartej na sieciach neuronowych”²³ (*neural network computing*) za podstawę prac nad widzeniem maszynowym wzorowanym na modelu ludzkiego mózgu. Ich zdaniem „informatyka oparta na sieciach neuronowych” była mylnym określeniem, gdyż u jej źródła leżała chybiona analogia między sieciami neuronowymi w mózgu a sieciami komputerowymi, niemająca wiele wspólnego z tym, jak biologia faktycznie pojmuje i bada struktury neuronowe.

21 Czyniąc to strategiczne rozróżnienie, pamiętam o myśli sformułowanej przez Humberta Maturanę i Francisca Varełę w *Autopoiesis and Cognition*. Zauważają oni, że postrzeganie już jest formą poznania, choć poznanie oznacza dla nich coś więcej niż czystą myśl. W ich ujęciu bowiem i postrzeganie, i poznanie są specyficznie biologicznymi zjawiskami. Zob. H.R. Maturana, F.J. Varela, *Autopoiesis and Cognition. The Realization of the Living*, Reidel, Dordrecht 1972.

22 G.N. Reeke Jr., G.M. Edelman, *Real Brains and Artificial Intelligence*, „Daedalus” 1988, nr 117 (1), s. 143.

23 Koncepcja sieci neuronowych została wprowadzona przez dwóch naukowców z University of Chicago, Warrena McCullocha i Waltera Pittsa, w roku 1944. Od tamtej pory sieci neuronowe wypadały z łask i powracały do łask w badaniach nad sztuczną inteligencją, lecz jednoznacznie umocniły swoją pozycję w najnowszych dokonaniach w dziedzinie maszynowego uczenia się.

Ściśle komputacyjne podejście, dowodzili, niewiele powie nam o percepcji, ponieważ zasadza się na błędzie, którym jest mniemanie, że

Przedmioty i wydarzenia, kategorie i logika są dane oraz że zadaniem mózgu jest przetwarzanie informacji o świecie przez algorytmy w celu wyciągnięcia wniosków prowadzących do pojawienia się zachowania [...]. Ów „problem kategorialny” bezpośrednio skutkuje niezdolnością systemów AI do radzenia sobie ze złożonością i nieprzewidywalnością rzeczywistego świata²⁴.

Można by przypuszczać, że zbudowane z warstw węzłów rekursywne sieci neuronowe stosowane w maszynowym, czy też głębokim, uczeniu się niwelują częściowo problem nieumiejętnego obchodzenia się z rosnącym skomplikowaniem i niepewnością świata. I owszem, udaje im się z powodzeniem identyfikować wzorce i trendy występujące w nieprecyzyjnych danych, co widać w aplikacjach takich jak rozpoznawanie twarzy, analiza danych medycznych, przewidywanie pogody czy tłumaczenie języków naturalnych. Jednakże sieci neuronowe w postaci obecnie przyjmowanej w badaniach nad sztuczną inteligencją nadal nie kwestionują założenia, „że informacja istnieje w świecie”, organizm zaś „jest o d b i o r c ą , a nie t w ó r c ą kryteriów prowadzących do informacji [wyróżnienie w oryginale]”²⁵. Szczególnie interesuje mnie przedstawiona przez Edelmana i Reeke’a krytyka wyobrażenia, że przedmioty i wydarzenia istnieją w świecie, niejako czekając na to, że je zobaczymy, pojmimy i wykorzystamy. Krytyka ta jest zbieżna z filozoficznym stanowiskiem znanym jako „świadomy realizm”, propagowanym przez psychologa kognitywnego Donalda D. Hoffmana²⁶. Nie chodzi o to, że w świecie niczego nie ma, tylko o to, że z góry założone, wyosobnione przedmioty, które rzekomo ukazują się naszemu wzrokowi, aby je uchwycił, są w istocie efektem twórczego, dynamicznego, a przy tym nieokreślonego procesu. Widzimy to, co potrzebujemy zobaczyć, twierdzi Hoffman. Jego zdaniem potrzeba ta ma charakter biologiczny – czy, dokładniej mówiąc, ewolucyjny – w tym sensie, że celem tego, co widzimy, jest nasze przetrwanie. W tym właśnie punkcie intelektualna trajektoria Hoffmana schodzi się z trajektorią Edelmana i Reeke’a.

24 G.N. Reeke Jr., G.M. Edelman, *Real Brains...*, s. 145.

25 Tamże, s. 153.

26 D.D. Hoffman, *The Case Against Reality. How Evolution Hid the Truth from Our Eyes*, MIT Press, Cambridge 2019.

W rzeczy samej w ripociście na obliczeniowy schematyzm fizyków Edelman i Reeke utrzymują, że wszelka sensowna teoria kategoryzacji i inteligencji użyteczna w badaniach nad sztuczną inteligencją musi przyjąć darwinowski model doboru, dostosowany jednak do działania neuronów pojedynczego organizmu funkcjonujących w trakcie jego życia. Wniosek ten posłużył im za punkt wyjścia w artykule z roku 1990, traktującym o możliwości skonstruowania maszyny percepcji.

I faktycznie w tekście *Is It Possible to Construct a Perception Machine?* Edelman i Reeke proponują taki konstrukt. Architektura tej maszyny, nazwanej, i to bez żadnej ironii, Darwinem III, zasada się na ich spostrzeżeniu, że kluczowa w układach percepcji jest kategoryzacja (a nie bardziej bezpośrednie rozpoznawanie obiektów, stosowane w systemach widzenia komputerowego). Osią postrzegania jest wprawdzie „przystosowawcze rozróżnianie przedmiotów lub wydarzeń przez jedną lub więcej modalności zmysłowych, wydzielające je z tła i od innych przedmiotów lub wydarzeń”²⁷, z tą jednak różnicą, że przedmioty owe nie są z góry określone. Według Edelmana i Reeke’a w percepcji nie chodzi o uchwycenie reprezentacji, gdyż kategorie, którymi posługuje się organizm lub maszyna, nie istnieją w świecie, czekając, aż ów organizm lub owa maszyna je po prostu rozpozna: kategorie te są aktywnie konstruowane (tj. znaczą coś dla postrzegającego „w danej sytuacji”²⁸). Percepcja zatem zyskuje nową definicję – „aktywnego procesu czuciowo-ruchowego wymagającego zgłębiania i zależącego od uprzednich doświadczeń”²⁹. W tym modelu kategoryzacja to czynność biologiczna, a nie matematyczna, gdyż nie da się jej wyrazić jedynie w postaci symbolicznej. Biologiczny wymiar organizmów wskazuje, że rozpoznanie, a więc również postrzeganie i inteligencja zachodzą w warunkach poruszania się w świecie. Podstawową przesłankę tego modelu Edelman i Reeke sformułowali w opublikowanym dwa lata wcześniej artykule, głosząc, że mózg to „wybiórczy system działający w czasie somatycznym”³⁰. Ta ciekawa myśl współgra z coraz bardziej popularnym podejściem fenomenologicznym do projektowania sztucznej inteligencji.

Koncepcję maszyny percepcji wypracowaną przez Edelmana i Reeke’a krępuje jednak akulturowość jej modelu, w którym siłą napędową operacji

27 G.M. Edelman, G.N. Reeke Jr., *Is It Possible to Construct a Perception Machine?*, s. 36.

28 Tamże, s. 37.

29 Tamże.

30 G.N. Reeke Jr., G.M. Edelman, *Real Brains and Artificial Intelligence*, s. 157.

maszynowych jest przede wszystkim dobór naturalny. Autorzy odrzucają kulturowy przekaz informacji, gdyż ich zdaniem nie jest on istotny dla ewolucyjnego rozwoju działających systemów postrzegania. Chociaż Darwin III ma być silnie osadzony w świecie, jego nastawienie, wizję i cele paraliżuje czasowość jego ruchu, jako że czas somatyczny wymaga wymazania wyjątkowości rzeczowej somy kosztem (ponoć) ponadczasowego funkcjonowania doboru naturalnego. Istnieje zatem niebezpieczeństwo, że maszyna percepcji Edelmana i Reeke'a skończy jako dość trywialna maszyna, gdyż nawet jeśli w celu wypracowania swoich kategorii będzie gromadzić informacje w dynamiczniejszy i bardziej zakotwiczony sposób, będą one z definicji wyzute z wszelkiej specyfiki kulturowej, która traktowana jest tu jedynie jako szum. Jednakże ustalenie, które aspekty otoczenia sklasyfikować można jako naturę, a które przynależą do domeny kultury, nie jest wcale oczywiste.

W znamienne zatytułowanym artykule *Beyond Biology and Culture The Meaning of Evolution in a Relational World* (Poza biologią i kulturą. Znaczenie ewolucji w relacyjnym świecie) antropolog Tim Ingold³¹ dowodzi, że różnica między naturą a kulturą jest kwestią jedynie czasowości: historia (która objawia się jako „kultura” w różnych epokach) działa w skalach czasowych o mniejszej rozpiętości niż ewolucja (która zachodzi według praw „natury”). W następnym kroku Ingold pokazuje, jak traktat *O pochodzeniu gatunków* Darwina podważył osiemnastowieczny obraz „człowieka” wyposażonego w szczególne cechy, takie jak rozum i moralność – walory, które miały człowieka odróżniać od innych gatunków. Natomiast Darwin utrzymywał, że różnica między człowiekiem a innymi zwierzętami jest zasadniczo różnicą stopnia, nie zaś rodzaju, oraz że ewolucja, którą pierwotnie nazywał „dziedziczeniem z modyfikacją”, to sposób, w jaki dokonała się zmiana prowadząca do racjonalnego człowieka z jego intelektualnymi i moralnymi przymiotami. Problematyczne jest to, że ów pogląd dał początek kolejnemu rozróżnieniu, które dziś nazwalibyśmy (słusznie zresztą) rasistowskim, a mianowicie rozróżnieniu między człowiekiem dzikim a cywilizowanym. Jedynym rozwiązaniem tego dylematu było przypisanie „historyczności procesowi kultury różniącemu się od procesu ewolucji biologicznej rodzajem, a nie stopniem” – wyjaśnia Ingold³². To z kolei wiązało się z ideą dwóch równoległych rodzajów dziedziczenia w ludzkich populacjach: dziedziczeniem biologicznym

31 T. Ingold, *Beyond Biology and Culture. The Meaning of Evolution in a Relational World*, „Social Anthropology” 2004, nr 2.

32 Tamże, s. 213.

(polegającym na przekazie informacji genetycznej zakodowanej w DNA i dotyczącym sedna tożsamości człowieka, np. chodzenia) oraz dziedziczeniem kulturowym (zachodzącym na drodze społecznego uczenia się, np. gry na instrumencie muzycznym). Przy czym chodzenie nie jest przecież „czysto” naturalne, a i nauka gry na instrumencie nie jest wyłącznie kulturowym doświadczeniem, odizolowanym od przekazu ucieleśnienia: „Specyficzne sposoby działania, postrzegania i poznania, które zwykliśmy nazywać kulturowymi, wplecione są w konstrukcję ludzkiego organizmu w przebiegu rozwoju ontogenicznego”, co czyni je w równym stopniu „faktami biologicznymi”³³. Ingold podkreśla, że problem nie leży właściwie w utożsamieniu biologii z genetyką. Dlatego też rozdzielenie procesów zmiany historycznej i biologicznej powinno się traktować jako odczasowienie (*detemporalisation*) – proces bazujący na fikcyjnym ustanowieniu oddzielnych porządków czasowych dla pewnych rodzajów zmian. Ten błędny model nadal obowiązuje i wyznacza obecny konsensus dotyczący metody naukowej oraz dyscyplinarny podział na nauki ścisłe i humanistyczne.

Widzenie maszynowe a epistemiczna (nie)sprawiedliwość

Wywód Ingolda kwestionuje wszelkie dążenia do skonstruowania systemów percepcji zdolnych do przekazania swoich cech ewolucyjnych potomkom, lecz niezdolnych do przekazania im wpływów kulturowych (czy też, w rzeczy samej, absolutnie wolnych od takich wpływów). Wszystkie cechy, które z o s t a n ą przekazane, będą nieodmiennie niosły niemożliwe do łatwego rozdzielenia inskrypcje „naturalne” i „kulturowe”, a konstrukcja kultury jako osobnej domeny nieodziedziczalnych cech będzie umacniać to rozróżnienie, pozwalając informatykom pomijać ucieleśnione i umiejscowione tryby postrzegania i poznania. Taki odcieleśniony model widzenia komputerowego utrwała jeden z największych mitów nauk ścisłych (i informatyki), a mianowicie przekonanie, że wyeliminowanie stronniczości danych rozumianej jako uprzedzenie kulturowe wystarczy do zapewnienia zarówno czystości, jak i sprawiedliwości danych³⁴. Raz po raz dowiadujemy się o konsekwencjach

33 Tamże, s. 216.

34 Sięgając po filozofię Bernarda Stieglera, Claudio Celis Bueno i María Jesús Schultz Abarca podkreślają, że „«wyłączanie» odziedziczonych uprzedzeń, aby ujrzeć rzeczywistość «samą w sobie», jest złudzeniem maskującym fakt, że technologia uporczywie modyfikuje nasze wewnętrzne zmysły postrzegania i pamięci. Nagi ludzki wzrok jest już z góry wzrokiem maszyno-

takiej esencjalizacji biologii i „mózgu” w kognitywistyce i informatyce kosztem „cech kulturowych”. Do znanych przykładów z roku 2020 należy algorytm zmiany tła w oprogramowaniu do wideokonferencji Zoom, który usuwał głowę czarnoskórego badacza, gdy tylko ten chciał zastosować wirtualne tło, oraz algorytm przycinania obrazu Twittera, który preferował białe twarze, ukazując przycięte obrazy na osi czasu.

Jak się okazuje, maszyna widzenia komputerowego nie jest szczególnie spostrzegawcza, a skutki jej rasistowsko nacechowanych defektów nie są bynajmniej trywialne. Algorytmy nią zawiadujące to przecież te same algorytmy, które podejmują decyzje dotyczące społecznego, finansowego i prawnego statusu jednostek i których efektem są sankcje podczas kontroli granicznej, odmowa kredytu, przewidywanie niepowodzenia szkolnego czy też przypisywanie przestępczości. W roku 2015 wytknięto firmie Google, że jej wczesny algorytm wykrywania twarzy automatycznie znakuje zdjęcia osób czarnoskórych znacznikiem „goryl”. Problemy z rozpoznawaniem twarzy czarnoskórych kobiet wciąż trwają, a wysoki odsetek błędnych rozpoznań branżowi eksperci tłumaczą trudnością w oświetlaniu czarnej twarzy z nałożonym makijażem. Wypierając się stronniczości, Thorsten Thies, dyrektor działu algorytmów niemieckiej firmy Cognitec, która dostarcza rządowi systemy rozpoznawania twarzy, wyjaśnił w niepokojąco rozbijającym sposób, że „trudniej jest zrobić dobre zdjęcie osobie o ciemnej skórze niż białej”³⁵. Decyduje o tym między innymi fakt, że bazy danych ze zbiorami zdjęć, które służą jako zestawy treningowe dla algorytmów tego typu, nie są reprezentatywne, gdyż pod względem zarówno ilości, jak i jakości przeważają w nich fotografie białych mężczyzn. Działa tutaj jednak również głębsza logika, za którą stoi cała systemowa infrastruktura związana z produkcją aparatów, systemów oświetleniowych, oprogramowania do obróbki zdjęć, a także wizualna i kulturowa edukacja fotografów i techników, która wytwarza zestaw zinternalizowanych norm przedstawianych następnie jako źródło „obiektywnych” utrudnień przy fotografowaniu osób o ciemnej skórze. Ten sposób myślenia, wrośnięty głęboko w rozmaite technologie poprzedzające cyfrowość, poddaje krytyce Safiya Noble w książce *Algorithms of Oppression* (Algorytmy ucisku). Przewiduje ona, że „w XXI wieku sztuczna inteligencja stanie się jednym

wym. Widzenie człowieka, tak jak widzenie maszyny, zależy od powierzchni inskrypcji funkcjonujących jako zewnętrzna władza wyobraźni”; C. Celis Bueno, M.J.S. Abarca, *Memo Akten's „Learning to See”*. From *Machine Vision to the Machinic Unconscious*, „AI & Society” 2020, s. 1185.

35 T. Simonite, *The Best Algorithms Struggle to Recognize Black Faces Equally*, „Wired” 22 lipca 2019.

z głównych problemów w obszarze praw człowieka³⁶, a tym samym wysłała sygnał ostrzegawczy i nawołuje, abyśmy dostrzegli, jaką rolę odgrywają architektura i logika algorytmicznych narzędzi decyzyjnych w kamuflowaniu i pogłębianiu niesprawiedliwości społecznej. Jej analiza kładzie kres przekonaniu, że krzywdzące decyzje systemowe to jedynie przypadkowe aberracje, które można łatwo usunąć, a dzięki temu przywrócić funkcjonalność i sprawność rzekomo neutralnego systemu.

Gdy w 2020 roku wybuchł skandal związany z algorytmem przycinania zdjęć, Dantley Davis, dyrektor do spraw danych Twittera, dość przewidywalnie bił się w pierś: „To w stu procentach nasza wina. [...] Teraz zajmiemy się rozwiązaniem tego problemu”³⁷. Lecz Noble nie pozostawia żadnych złudzeń, podkreślając, że „algorytmiczna opresja” nie jest jedynie błędem systemu, jest ona raczej zasadniczym elementem systemu operacyjnego sieci³⁸. Mitra Azar, Geoff Cox i Leonardo Impett również zauważają, że

W społeczeństwie o strukturalnej nierówności niezwykle trudno jest stworzyć „sprawiedliwy” algorytm; a już całkowicie niemożliwe jest stworzenie algorytmu, który byłby zarówno sprawiedliwy, jak i efektywny. [...] W niesprawiedliwym społeczeństwie maszyna klasyfikująca zawsze będzie niesprawiedliwa (przynajmniej w jednym sensie)³⁹.

Nie wystarczy zatem uporać się ze stroniczością danych. Zamiast tego należy postawić szersze pytania o formy niesprawiedliwości tkwiące w samych systemach je gromadzących i udostępniających. Musimy też dociekać, co to właściwie oznacza, że wyeliminowanie błędu, choć pożądane z technicznego punktu widzenia, koniec końców tym bardziej usprawnia sankcjonujący nadzór napędzany

36 S. Noble, *Algorithms of Oppression*, New York University Press, New York 2018, e-book.

37 @dantley na Twitterze, 19 września 2020.

38 S. Noble, *Algorithms of Oppression*.

39 M. Azar, G. Cox, L. Impett, *Introduction. Ways of Machine Seeing*, „AI and Society” 2021, t. 36, s. 1103. Numer specjalny czasopisma „AI and Society” zatytułowany *Ways of Machine Seeing* (Sposoby widzenia maszynowego) poświęcony jest wielodyscyplinarnym analizom zagadnień widzenia maszynowego i związanych z nim problemów. Przyjmując za punkt wyjścia sformułowane przez Johna Bergera w jego kanonicznych *Sposobach widzenia* stwierdzenie, że „każdy obraz ucieleśnia jakiś sposób widzenia” (J. Berger, *Sposoby widzenia*, przeł. M. Bryl, Aletheia, Warszawa 2023, s. 10), autorzy publikujący w numerze dociekają, jak „maszyny, a szczególnie technologie obliczeniowe, zmieniają sposób, w jaki widzimy świat” (M. Azar, G. Cox, L. Impett, *Introduction*, s. 1093).

przez te dane. Korekta stronniczości danych nie koryguje brutalnie penetrującej i ekstraktywistycznej logiki systemu widzenia komputerowego, a wręcz przeciwnie – umacnia ją. Można zatem stwierdzić, że aby w ogóle zacząć rozważać budowę nietrywialnej maszyny *p e r c e p c j i*, musimy najpierw rozpoznać niesprawiedliwe działania nietrywialnej maszyny *w i d z e n i a*.

W ramach projektu „inżynierii pojęciowej” nakreślonego na wstępie zmierzamy zatem do koncepcyjnego poszerzenia pojęcia maszyny. Zgodnie z cybernetycznym dziedzictwem sięgającym von Foerстера i innych badaczy termin „maszyna” odchodzi od ściśle inżynierskich skojarzeń i zaczyna obejmować układy wszelkiego rodzaju, czy to mechaniczne, czy biologiczne, o rozmaitych stopniach złożoności. W takim ujęciu systemy – od mikroskopijnych po kosmiczne – mieszczą się jedne w drugich i przekształcają się wewnątrznie, a nawet podlegają międzysystemowym przemianom. W latach sześćdziesiątych i siedemdziesiątych XX wieku cybernetyczne pojęcie maszyny stało się nośną koncepcją, po którą sięgali filozofowie i teoretycy kultury, aby wyartykułować różne poziomy złożoności społeczno-politycznej przy uwzględnieniu biologicznego i technicznego konstytuowania się zarówno poszczególnych jednostek, jak i społeczeństwa jako całości. Dobrej ilustracji tego zjawiska dostarczają maszyny wojenne i maszyny pragnące Gilles’a Deleuze’a i Félix’a Guattariego⁴⁰, *dispositif*⁴¹ Michela Foucaulta⁴² czy aparat Viléma Flussera⁴³. Czerpiąc z tej spuścizny, powinniśmy założyć, że „maszyna percepcji”, którą staramy się zbudować, obejmie wielorakie, acz splecione, poziomy znaczeniowe. Koncepcja ta uwyrażnia zazębianie się dyskursów naukowych i kulturowych w wytwarzaniu obrazów oraz fakt, że subiektywizm i obiektywizm kształtują się jako funkcje wszystkich rodzajów obrazów.

W mojej próbie zbudowania nietrywialnej maszyny percepcji nie chodzi zatem o skonstruowanie jakiegoś urządzenia, a przynajmniej nie bezpośrednio

40 G. Deleuze, F. Guattari, *Tysiąc plateau. Kapitalizm i schizofrenia*, t. 2, przeł. J. Bednarek, M. Herer, Bęc Zmiana, Warszawa 2015; ciż, *Anty-Edyp. Kapitalizm i schizofrenia*, przeł. T. Kaszubski, Wydawnictwo Krytyki Politycznej, Warszawa 2017.

41 W polskich tłumaczeniach pism Michela Foucaulta proponowane są różne odpowiedniki tego pojęcia, do najczęstszych z nich należą: „dyspozytyw”, „urządzenie” i „aparat”. Zob. np. B. Nowicka, „Urządzenie”, „zastosowanie”, „układ” – *kategoria dispositif u Michela Foucaulta, jej tłumaczenia i ich implikacje dla postfoucaultowskich analiz władzy*, „Przegląd Socjologii Jakościowej” 2011, nr 2 (7) [przyp. tłum.].

42 M. Foucault, *The Confessions of the Flesh, w: tegoż, Power/Knowledge. Selected Interviews and Other Writings*, red. C. Gordon, Pantheon Books, New York 1980.

43 V. Flusser, *Ku filozofii fotografii*; tenże, *Into the Universe of Technical Images*.

o to; chodzi raczej o zbudowanie ramy pojęciowej służącej zarówno przedstawieniu tego, co nazywamy „światem”, jak i widzeniu tego świata lepiej. Taka maszyna percepcji będzie miała do wykonania więcej zadań niż tylko rozpoznawanie symboli, unikanie uprzedzeń czy nawet liczenie się ze środowiskiem. W jej architekturze idea mózgu jako wyodrębnionego narządu percepcji będzie musiała ustąpić miejsca dynamicznej interakcji między organizmem, z jego nieustannie zmieniającym się ucieleśnieniem, a środowiskiem, które nie jest stałą, lecz „istnienie wyłączone w relacji z zamieszkującymi je organizmami i ucieleśnia historię współoddziaływania z nimi”⁴⁴. W rzeczy samej maszyna percepcji nie istnieje poza swoim kulturowym i historycznym umiejscowieniem. Sam gest podjęcia starań, aby zbudować nietrywialną maszynę percepcji, ma również stanowić wielopoziomą próbę przemyślenia na nowo naszego ludzkiego postrzegania i widzenia, a także zakwestionowania parametrów maszyny widzenia kształtującej się w naszym globalnie usieciowionym świecie. (Stanowisko to pokrótce wyraża założenia mojego projektu inżynierii pojęciowej).

Parametry maszyny widzenia przenikliwie przeanalizował Paul Virilio w *Maszynie widzenia* opublikowanej w języku francuskim w roku 1988 (*La machine de vision*), a w języku angielskim w roku 1994 (*Vision Machine*)⁴⁵. Virilio odnajduje strukturalne początki tej maszyny w przekształceniach widzenia w XX wieku: od militarnych technologii pierwszej wojny światowej, które obróciły widzenie w wizualizację, aż po uprzemysłowienie widzenia, w wyniku którego większość czynności wykonywanych niegdyś przez ludzi w polu wizualnym powierzono maszynom. Virilio twierdzi również, że nadanie widzeniu charakteru przede wszystkim maszynowego szło w parze z „automatyzacją postrzegania”⁴⁶ spowodowaną postępami sztucznej inteligencji. Postuluje jednak wyraźny zwrot od widzenia bardziej organicznego (i, z pozoru, przedtechnicznego) ku jego maszynowemu odpowiednikowi, o czym pisze następująco: „relatywistyczna fuzja/konfuzja tego, co faktyczne (albo – jeśli kto woli – operacyjne), i tego, co wirtualne; wyższość «efektu rzeczywistości» nad zasadą rzeczywistości”⁴⁷. Zwrot ten, zapoczątkowany wynalezieniem protez wzroku, takich jak teleskop

44 T. Ingold, *Beyond Biology and Culture*, s. 218.

45 P. Virilio, *Vision Machine*, Indiana University Press, Bloomington 1994. Polski przekład ostatniego rozdziału tej książki znajduje się w tomie *Widzieć, myśleć, być. Technologie mediów* pod redakcją A. Gwóźdźa. Zob. P. Virilio, *Maszyna widzenia*, przeł. B. Kita, w: *Widzieć, myśleć, być. Technologie mediów*, wybór, wstęp i oprac. A. Gwóźdź, Universitas, Kraków 2001 [przyp. tłum.].

46 P. Virilio, *Maszyna widzenia*, s. 39.

47 Tamże, s. 41.

w roku 1608, a nabierający tempa trzy stulecia później wraz z mnożeniem się „maszyn widzących”, Virilio interpretuje zasadniczo w kategoriach zamętu („konfuzji”) i straty. Wprawdzie polityczna krytyka zautomatyzowania percepcji jest wielce pożądana – a Virilio niewątpliwie takową podejmuje, przyglądając się wizualnym technologiom wojennym, automatyzacji widzenia w propagandzie i marketingu oraz mechanizacji sprawiedliwości w postępowaniach sądowych wykorzystujących wideo – lecz przedstawiona przez niego ontologiczna krytyka transformacji widzenia koniec końców utrwała wyobrażenie, że widzenie to przymiot ludzki i humanistyczny, oraz sytuje podmiotowość człowieka na płaszczyźnie metafizycznej. W efekcie zaś jego maszyna widzenia jest „autonomicznym układem technicznym”⁴⁸, który zaburza pierwotną organiczną jedność i czystość ludzkiego wzroku. Korzystając z niektórych idei Virilia, przyjmuję w tym artykule pojęcie maszyny złożonej z komponentów zarówno organicznych, jak i nieorganicznych, zmieniającej się w czasie, a jednocześnie na rozmaite sposoby kształtującej ludzkie sensorium. Prezentowana tu próba zbudowania nietrywialnej maszyny percepcji jest zbieżna z przeprowadzoną przez Bernarda Stieglera analizą automatyzacji w kategoriach farmakologii⁴⁹. Z analizy Stieglera wynika możliwość przepracowania tego procesu, ogólnie szkodliwego dla życia jednostki i społeczeństwa – czego dowodzą wszechobecny „złotań”, „nihilizm” i „wątpienie ludzkości we własną przyszłość”⁵⁰ – tak aby wyzwolić lecznicze właściwości maszyny.

48 J. Armitage, *Accelerated Aesthetics: Paul Virilio's „The Vision Machine”, „Angelaki”* 1997, nr 2 (3), s. 203.

49 Farmakologia to koncepcja wywodząca się z myśli Jacques'a Derridy (zob. tenże, *Farmakon*, przeł. K. Matuszewski, w: *Pismo filozofii*, wybrał i przedmową opatrzył B. Banasiak, inter esse, Kraków 1992). Stiegler sięgnął po nią w swoich późniejszych pracach poświęconych analizie uwarunkowań kapitalizmu obliczeniowego oraz społeczno-gospodarczej niedoli, której przysparza ta szczególna formacja polityczna. Stiegler, co ważne, nie ograniczał się jedynie do diagnozowania, lecz z zapałem poszukiwał sposobności w – i dróg wyjścia z – desperacji i nihilizmu sprowadzanego przez zautomatyzowaną decyzyjność i odindywidualizowanie człowieka w takich warunkach politycznych i ekonomicznych. Jak ujął to w *Automatic Society*, t. 1: „z takiej perspektywy, do kwestii innowacji należy podejść bardzo poważnie, a nie traktować jej li i tylko jako ideologicznego dyskursu opartego na marketingowych opowieściach. Innowacja ma ewidentnie zupełnie realną funkcję gospodarczą, gdyż bez wątplenia ustanawia wytwarzanie negentropii. Jest dziś dla nas oczywiste, że wytwarzana krótkoterminowo negentropia generuje długofalową rozleglejszą entropię. Cała kwestia organologii i jej farmakologii w obszarze negentropicznym tkwi w tym, że farmakon może być toksyczny i leczniczy jedynie w zakresie (i w nadmiarze), który jest zarówno entropiczny, jak i negentropiczny”; B. Stiegler, *Automatic Society*, t. 1: *The Future of Work*, przeł. D. Ross, Polity Press, Cambridge 2017, s. 100.

50 Tamże, s. 9, 7.

Maszyna widzenia, w której obecnie tkwimy, działa i jako aparatura nadzoru, i jako przyrząd do przechwytywania danych. Jest ona również, jak podkreśla Virilio, maszyną wojenną, służącą za cyfrowe pole bitwy, na którym toczą się mnogie operacje przechwyty oraz przemocy, mające rzeczywiste konsekwencje dla istnień ludzkich i nie-ludzkich. Nietrywialna maszyna percepcji, którą chciałabym zbudować, musiałaby być zdolna przebić się przez jej mętną logikę. Taka zdolność z kolej wymagałaby, aby maszyna percepcji wyszła poza neutralność i niestronniczość (choć powinna być neutralna i niestronnicza), a mianowicie, aby była zdecydowanie antyrasistowska. Moglibyśmy posunąć się dalej i zasugerować, że aby przeciwstawić się i rasistowskiemu dziedzictwu maszyny wojennej, i kapitalistycznemu ekstraktywizmowi, który ją napędza, maszyna percepcji musiałaby również być antyimperialistyczna, jak dowodzi Ariella Aïsha Azoulay. W książce *Potential History. Unlearning Imperialism* (Potencjalna historia. Oduczanie się imperializmu) Azoulay uderza w układ praw i przywilejów, który od początku decydował o mechanicznym wytwarzaniu obrazów i orzekł, że „świat jest po to, aby go pokazywać”, i to „tylko wybranym widzom”⁵¹. Azoulay twierdzi wręcz, że „fotografia rozwinęła się wraz z imperializmem; aparat uwidoczniał niszczenie świata i uczynił je akceptowalnym oraz legitymizował odbudowę świata na warunkach imperialnych”⁵².

Na pierwszy rzut oka mogłoby się wydawać, że wywód Azoulay stawia zarzuty tylko fotografii humanocentrycznej, która z założenia ma być pokazywana (wybranym) ludziom i przez nich widziana, lecz niekoniecznie zwraca się przeciwko wielkim bazom danych nadającym dziś kształt widzeniu maszynowemu, któremu dostarczają pożywki. Lecz przecież ta sama forma racjonalności leży u podstaw wytwarzania, przechowywania i kategoryzowania wszystkich mechanicznie produkowanych obrazów. Ich konstytutywna logika, historia i sposoby przedstawiania niezmiennie kierują się imperialnym myśleniem, który legitymizowało klasyfikację jako ponoć neutralną, a jednocześnie zaprzęgało ją do działań podporządkowanych celom Imperium. Imperialne rządy mają dziś inne umiejscowienie i inny format niż niegdyś: w znamiennie zatytułowanej książce Michael Hardt i Antonio Negri dowodzą, że Imperium nie ma już ani konkretnych granic, ani terytorialnie umocowanego ośrodka władzy. Stało się bowiem „*d e c e n t r a l i z o w a n y m*

51 A.A. Azoulay, *Potential History. Unlearning Imperialism*, Verso, London 2019, s. 4.

52 Tamże, s. 5.

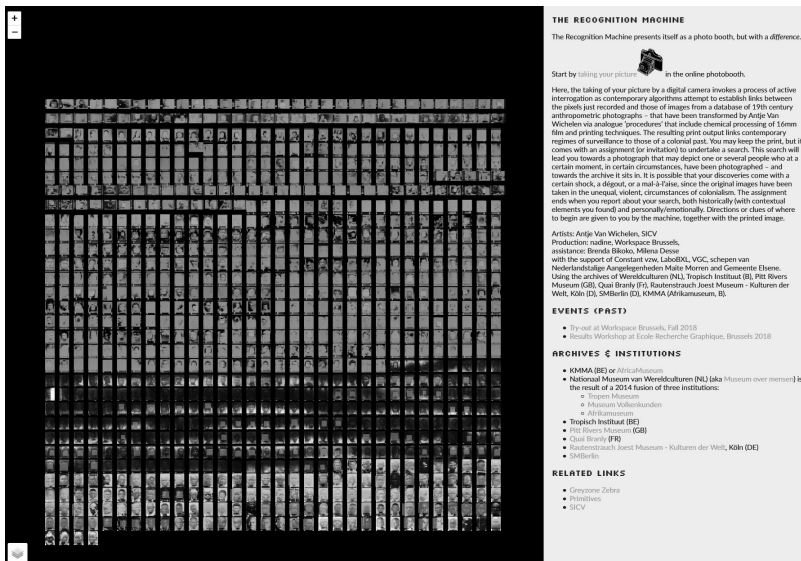
i a terytorialnym aparatem rządzenia [wyróżnienie w oryginale]⁵³, a globalne przepływy kapitału realizują pewną formę biopolityki, uczestnicząc w wytwarzaniu życia społecznego jako takiego, które ma postać węzła przenikających się sił ekonomicznych, politycznych i kulturowych. Neoimperialna maszyna wojenna jest zatem nade wszystko hegemoniczną maszyną nadzoru: dokonuje podbojów za pomocą domniemanej zgody oraz dzięki skali i niewidoczności swej ingerencji. Nie oznacza to, że w nowym aparacie imperialnym nie sposób zlokalizować punktów węzłowych, w których skupiona jest władza. Jednakże coraz powszechniejsze i dalej sięgające przeniesienie dominacji i decyzyjności z gestii rządów w gestię korporacji, z Waszyngtonu do Doliny Krzemowej, z USA do Chin i z ludzi na algorytmy kreuje rozchwianą i rozmytą geopolitykę, w której polityczne i technologiczne czarne skrzynki maskują umiejscowienie władzy, a także jej faktyczne poczynania. Zarazem ciężar odpowiedzialności spada na ramiona poszczególnych obywateli, których główną, choć od razu podwójną, tożsamością w tym systemie jest tożsamość użytkowników Internetu, a jednocześnie punktów danych. Neoimperialna maszyna nadzoru stanowi więc zaktualizowaną wersję maszyny widzenia, którą opisywał Virilio. Nietrywialną maszynę percepcji można by natomiast określić jako jej pojęciowy i techniczny odpowiednik, dający jednak sposobność nowego widzenia nas samych, jak również świata.

Maszyna rozpoznawania

Za przykład możliwej realizacji nietrywialnej, a przy tym aktywnie antyimperialistycznej maszyny może posłużyć *The Recognition Machine* („Maszyna rozpoznawania”) stworzona przez artystów Antje Van Wichelen i SICV (Michael Murtaugh i Nicolas Malevé) (il. 3 i 4). Miałam okazję uczestniczyć w festiwalowej odsłonie tej pracy podczas Photoszene w Kolonii w maju 2019 roku, ale projekt ma również wersję internetową. *The Recognition Machine* wygląda jak budka fotograficzna, do której widzowie mogą wejść i zrobić sobie zdjęcie cyfrowe. Akt fotografowania się aktywuje algorytm, który próbuje „ustanowić powiązania między właśnie zarejestrowanymi pikselami a pikselami zdjęć z bazy danych zawierającej dziewiętnastowieczne fotografie antropomorficzne”, obrobione technikami analogowymi. „Powstające w ten sposób wydruki łączą współczesne systemy nadzoru z tymi z kolonialnej przeszłości”⁵⁴.

53 M. Hardt, A. Negri, *Imperium*, przeł. S. Ślusarski, A. Kołbaniuk, W.A.B, Warszawa 2005, s. 8.

54 Zob. *The Recognition Machine*, b.d., <https://recognitionmachine.vandalist/> (31.01.2025).



Ilustracje 3 (ta strona) i 4 (następna strona). Antje Van Wichelen i SICV, *The Recognition Machine*, 2018.

Osią łączącą zdjęcia są emocje rozpoznawane przez algorytm na twarzy widza i skojarzone z emocjami odczytanymi z archiwalnych fotografii. Odczytania te uzyskano, trenując algorytm przy użyciu zbioru danych FER-2013, w którym każde zdjęcie przyporządkowano jednej z siedmiu emocji: gniewowi, wstrętowi, strachowi, szczęściu, smutkowi, zaskoczeniu i odczuciom neutralnym. Każde błędne rozpoznanie emocji zachodzące w tym procesie to ostrzeżenie sygnalizujące konsekwencje, którymi najeżona jest technologia predykcyjna: w tym przypadku znakowanie jest jedynie niewinną zabawą publiczności na wystawie, lecz poza galerią mylne rozpoznanie powiązań między obrazami, ich błędna kategoryzacja i niewłaściwe przypisanie niosą poważne skutki rzutujące na życie wielu ludzi. Widzowie mogą zachować otrzymane wydruki, zachęca się ich także do namysłu nad domniemaną analogią, a zatem do zastanowienia nad archiwum i kolonialną historią portretowania. *The Recognition Machine* uświadamia nam również, że wszystkie obrazy funkcjonują jako część imperialno-kolonialnej sieci wizualności, która pewne ciała uznaje za widzialne i właściwe, inne zaś za nieczytelne i/lub nielegalne. Interującym elementem tego projektu jest fakt, że twórcy odchodzą od idei pojedynczego obrazu będącego artefaktem podziwianym, klasyfikowanym i wykorzystywanym na inne sposoby w odosobnieniu od innych: pokazują



oni, że w s z y s t k i e obrazy są częścią wielorakich sieci wymiany wiedzy i danych. *The Recognition Machine* przedstawia model maszyny percepcji jako zaproszenie do zgłębiania procesów wytwarzania wizualności, sieci obrazów i ich infrastruktury, danych i baz danych, które leżą u ich podstaw, oraz algorytmów, które kształtują zarówno ich wytwarzanie, jak i ich wzajemne sieciowe powiązania. Można zatem stwierdzić, że n i e t r y w i a l n a maszyna percepcji musiałaby skłaniać do zaangażowanego, etyczno-politycznego podejścia do obrazów, ich historii, baz danych i infrastruktury. Powinna również obejmować strategie umożliwiające ludziom korzystanie z baz danych w celu od-przemysłowienia wizualności i widzenia. Chodzi zatem nie tylko o to, aby zobaczyć, co tkwi wewnątrz maszyny-archiwum i jak ona myśli, ale też o stworzenie warunków do myślenia inaczej o widzeniu ludzkim i maszynowym. Wprawdzie nie można oczekiwać od ludzi zdolności ujrzenia wszystkich obrazów dostępnych w licznych bazach i chmurach danych, przesłedzenia wszystkich możliwych powiązań między nimi czy poznanie wszystkich oferowanych kategorii i znaczników. Lecz ludzie-widzowie mają zdolność, a wręcz obowiązek, wziąć pod lupę logikę nieprzejrzystości i skali, która zawiaduje napędzaną przez sztuczną inteligencję maszyną percepcji, zapytywać o jej działania, a także domagać się lepszego – sprawiedliwszego,

bardziej historycznie osadzonego oraz wyraźniej antyrasistowskiego i anty-imperialistycznego – podejścia do przepływu obrazów i danych. Aby odczytać się imperializmu, jak dowodzi Azoulay, musimy przyjrzeć się „konceptyjnym źródłom imperialnej przemocy – przemocy, która zakłada, że ludzie i światy to surowiec z góry przynależny do imperialnych zasobów”⁵⁵. Rozmyte granice dzisiejszego Imperium w powiązaniu z korzyściami, które jednostki czerpią ze stania-się-danymi dla serwotechnologii kapitalizmu platform, sprzyjają przyzwoleniu dla tej formy przemocy, a nawet sprawiają, że w ogóle nie uchodzi ona za przemoc. Jeśli poczynimy pierwsze kroki ku zbudowaniu nietrywialnej maszyny percepcji – która ma posłużyć jako matryca alternatywnej wersji społeczeństwa i jego sposobu ujmowania świata – zobaczymy wyraźnie, czym jest ta forma ekstraktywnej biopolitycznej przemocy i łatwiej będzie zaplanować, jak się jej przeciwstawić.

Przełożyła Patrycja Poniatowska

55 A.A. Azoulay, *Potential History*, s. 8.

Abstract

Joanna Żylińska

KING'S COLLEGE LONDON

Beyond Machine Vision: How to Build a Non-Trivial Perception Machine

Analyzing artificial intelligence and creativity through the lens of machine vision, this article examines the impact of the vision model developed in computer science on our social and political values and institutions. It proposes a creative "conceptual engineering" experiment to construct a non-trivial perception machine, drawing on two scientific articles about human-machine relations: Heinz von Foerster's "Perception of the Future and the Future of Perception," in which the concept of a "non-trivial machine" first appeared, and Gerald M. Edelman and George N. Reeke's "Is It Possible to Construct a Perception Machine?" Through a critical reading of these works, the author establishes conceptual foundations for a theory and praxis of machine perception, and reflects on the broader problem of epistemic and racial (in)justice in the technical industry concerned with making machines "see."

Keywords

machine vision, computer vision, perception, imperialism, bias